

# IOWA STATE UNIVERSITY

## Digital Repository

---

Animal Science Publications

Animal Science

---

9-10-2020

## Fecal and Ruminant Microbiome Components Associated With Methane Emission in Beef Cattle

Bruno G. N. Andrade  
*Embrapa Pecuaria Sudeste*

Haithem Afi  
*Cork Institute of Technology*

Flavia A. Bressani  
*Embrapa Pecuaria Sudeste*

Rafael R. C. Cuadrat  
*Dife*

Priscila S. N. de Oliveira  
*Embrapa Pecuaria Sudeste*

*See next page for additional authors*

Follow this and additional works at: [https://lib.dr.iastate.edu/ans\\_pubs](https://lib.dr.iastate.edu/ans_pubs)



Part of the [Agriculture Commons](#), [Animal Sciences Commons](#), [Bacteriology Commons](#), [Environmental Microbiology and Microbial Ecology Commons](#), and the [Statistical Models Commons](#)

The complete bibliographic information for this item can be found at [https://lib.dr.iastate.edu/ans\\_pubs/559](https://lib.dr.iastate.edu/ans_pubs/559). For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

---

This Article is brought to you for free and open access by the Animal Science at Iowa State University Digital Repository. It has been accepted for inclusion in Animal Science Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

---

# Fecal and Ruminal Microbiome Components Associated With Methane Emission in Beef Cattle

## Abstract

**Background:** The impact of extreme changes in weather patterns in the economy and humanity welfare are some of the biggest challenges that our civilization is facing. From the anthropogenic activities that contribute to climate change, reducing the impact of farming activities is a priority, since its responsible for up to 18% of greenhouse gases linked to such activities. To this end, we tested if the ruminal and fecal microbiomes components of 52 Brazilian Nelore bulls, belonging to two experimental groups based on the feed intervention, conventional (A) and byproducts based diet (B), could be used as biomarkers for methane (CH<sub>4</sub>) emission.

**Results:** We identified a total of 5,693 Amplicon Sequence Variants (ASVs) in the Nelore bulls microbiomes from the experimental group B. Statistical analysis showed that the microbiome populations were significantly different among treatment groups. Differential abundance (DA) analysis with the ANCOM approach identified 30 bacterial and 15 archaea ASVs as DA among treatment groups. Random forest models, using either bacteria or archaea ASVs as predictors, were able to predict the treatment group with high accuracy ( $r^2 > 0.85$ ). Association analysis using Mixed Linear Models indicate that bacterial and archaea ASVs are linked to the CH<sub>4</sub> emission phenotype, of which the most prominent were the ruminal ASV 40 and fecal ASV 35. These ASVs contributed to a 9.7% increase and 7.3% decrease of the variation in CH<sub>4</sub> emission, respectively, which indicated their potential as targets for feed interventions and/or biomarkers.

**Conclusion:** The feed composition induced significant differences in abundance and richness of ruminal and fecal microbial populations. The dietary treatment based on industrial byproducts applied had an impact on the microbiome diversity of bacteria and archaea, but not on protozoa. Microbiome components (ASVs) of bacteria and archaea can be successfully used to predict the treatment group, thus giving support to the hypothesis that the feed intervention modulate microbiome abundance and diversity. Microbiome components were associated with CH<sub>4</sub> emission in both microbiomes. Therefore, both ruminal and fecal ASVs can be used as biomarkers for methane production and emission.

## Keywords

Association, archaea, bacteria, Bos Indicus, Biomarkers, Random Forests

## Disciplines

Agriculture | Animal Sciences | Bacteriology | Environmental Microbiology and Microbial Ecology | Statistical Models

## Comments

This preprint is made available through Research Square, doi: [10.21203/rs.3.rs-74103/v1](https://doi.org/10.21203/rs.3.rs-74103/v1).

## Creative Commons License



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

## Authors

Bruno G. N. Andrade, Haithem Afi, Flavia A. Bressani, Rafael R. C. Cuadrat, Priscila S. N. de Oliveira, Gerson B. Mourão, Luiz L. Coutinho, James M. Reecy, James E. Koltes, Marcela Maria de Souza, Adhemar

---

Zerlotini Neto, Sérgio Raposo de Medeiros, Alexandre Berndt, Julio C. P. Palhares, and Luciana C. A. Regitano

# Fecal and Ruminant Microbiome Components Associated With Methane Emission in Beef Cattle.

**Bruno G.N. Andrade**

Embrapa Pecuaria Sudeste

**Haithem Afli**

Cork Institute of Technology

**Flavia A. Bressani**

Embrapa Pecuaria Sudeste

**Rafael R. C. Cuadrat**

Dife

**Priscila S. N. de Oliveira**

Embrapa Pecuaria Sudeste

**Gerson B. Mourão**

Universidade de Sao Paulo Escola Superior de Agricultura Luiz de Queiroz

**Luiz L. Coutinho**

Universidade de Sao Paulo Escola Superior de Agricultura Luiz de Queiroz

**James M. Reecy**

Iowa State University

**James E. Koltes**

Iowa State University

**Marcela Maria de Souza**

Iowa State University

**Adhemar Zerlotini Neto**

Embrapa Informatica Agropecuaria

**Sérgio Raposo de Medeiros**

Embrapa Pecuaria Sudeste

**Alexandre Berndt**

Embrapa Pecuaria Sudeste

**Julio C. P. Palhares**

Empresa Brasileira de Pesquisa Agropecuaria

**Luciana C. A. Regitano** (✉ [luciana.regitano@embrapa.br](mailto:luciana.regitano@embrapa.br))

Embrapa Pecuaria Sudeste <https://orcid.org/0000-0001-9242-8351>

**Keywords:** Association, archaea, bacteria, Bos Indicus, Biomarkers, Random Forests

**DOI:** <https://doi.org/10.21203/rs.3.rs-74103/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** The impact of extreme changes in weather patterns in the economy and humanity welfare are some of the biggest challenges that our civilization is facing. From the anthropogenic activities that contribute to climate change, reducing the impact of farming activities is a priority, since its responsible for up to 18% of greenhouse gases linked to such activities. To this end, we tested if the ruminal and fecal microbiomes components of 52 Brazilian Nelore bulls, belonging to two experimental groups based on the feed intervention, conventional (A) and byproducts based diet (B), could be used as biomarkers for methane (CH<sub>4</sub>) emission.

**Results:** We identified a total of 5,693 Amplicon Sequence Variants (ASVs) in the Nelore bulls microbiomes from the experimental group B. Statistical analysis showed that the microbiome populations were significantly different among treatment groups. Differential abundance (DA) analysis with the ANCOM approach identified 30 bacterial and 15 archaea ASVs as DA among treatment groups. Random forest models, using either bacteria or archaea ASVs as predictors, were able to predict the treatment group with high accuracy ( $r^2 > 0.85$ ). Association analysis using Mixed Linear Models indicate that bacterial and archaea ASVs are linked to the CH<sub>4</sub> emission phenotype, of which the most prominent were the ruminal ASV 40 and fecal ASV 35. These ASVs contributed to a 9.7% increase and 7.3% decrease of the variation in CH<sub>4</sub> emission, respectively, which indicated their potential as targets for feed interventions and/or biomarkers.

**Conclusion:** The feed composition induced significant differences in abundance and richness of ruminal and fecal microbial populations. The dietary treatment based on industrial byproducts applied had an impact on the microbiome diversity of bacteria and archaea, but not on protozoa. Microbiome components (ASVs) of bacteria and archaea can be successfully used to predict the treatment group, thus giving support to the hypothesis that the feed intervention modulate microbiome abundance and diversity. Microbiome components were associated with CH<sub>4</sub> emission in both microbiomes. Therefore, both ruminal and fecal ASVs can be used as biomarkers for methane production and emission.

## Background

Climate change caused by human activity is one of the biggest threats to our civilization [1, 2]. To mitigate its effects and sustain the feeding needs of an ever-growing human population, the efficient production of food, such as crops and animal farming, is a top priority [3]. Cattle farming, a valuable source of animal protein, is responsible for the emission of up to 18% of greenhouse gases of anthropogenic origin [4, 5], such as methane (CH<sub>4</sub>), a greenhouse gas 28 times stronger than CO<sub>2</sub>, therefore having a significant environmental impact. Although studies seeking to increase productivity and mitigate the environmental impact of cattle have been published over the years [6–8], only recently, the microbiome has started to be considered as an important subject for such studies [9, 10].

The term microbiome describes the total genomic content of populations of microorganisms in a given environment, also known as microbiota. These microorganisms can shape the host biology through beneficial interactions, and thereby influence health, development and immune system modulation [11, 12]. Some studies suggest that the microbiota profile has a genetic component [13]. However, most of these microorganisms are elusive and utterly unknown to science due to inherent difficulties related to lab procedures for cultivation [14]. However, thanks to the development of new sequencing technologies, we are now able to access their genetic material directly and investigate their identity, distribution, relatedness, and functionality using approaches from the meta'omics discipline, such as deep sequencing metagenomics, metabarcoding and metatranscriptomics [15].

This field has been an active object in areas related to human health [16–18] and biotechnology, having a direct impact on industries and food production, like agriculture [19] and animal farming [20, 21]. The relationship between the microbiome and animals is being explored in studies to identify patterns that could increase their efficiency, reduce costs and their environmental impact [10, 22]. However, to overcome differences in sequencing library sizes in metabarcoding studies, the data have to be grouped in fractions (frequencies) to be compared between samples. Metabarcoding (16S and 18S rRNA gene sequences) data is compositional and resides in a simplex rather than the Euclidean space [23], due to the sum constraint (frequencies of a sample sum to 1), and thus have to be investigated using approaches developed by the Compositional Data Analysis (CoDA) discipline. Researchers proposed data transformations approaches to remove the unit-sum constraint of compositional data, such as centered log-ratio transformation (CLR), additive log-ratio transformation (ALR), and isometric log-ratio transformation (ILR), of which CLR is most often used in multivariate data analysis [24, 25].

Recently, the microbiome structure of the Nelore Brazilian beef cattle breed has gained the attention of the scientific community, being investigated in different studies [26, 27], including a previous study by our research group [28], in which we investigated the taxonomic profile of 26 Nelore bulls microbiomes and the co-occurrence of ruminal and fecal ASVs. Herein, we extended that study by the introduction of an additional experimental group under a different dietary intervention and compared the microbiome populations from two distant sections of the Nelore gastrointestinal tract (GIT), rumen and rectal ampulla in order to: (i) Describe the microbiome structure of animals fed with industrial byproducts; (ii) identify the impact of the dietary treatment on the microbiome diversity and abundance; (iii) identify associations between microbiome components and CH<sub>4</sub> emission under two dietary treatments.

## Methods

### Experimental design, sample collection and processing

The experimental population consisted of animals born in 2014, slaughtered in 2016, and divided into two groups based on the dietary treatment. The first experimental group (Group A, n=26), consisted of animals fed with a conventional diet based on corn silage, corn and soybean meals as concentrate as described in our previous study [28]. The second experimental group had a total replacement of

concentrates with the industrial by-products citrus pulp, corn germ, corn germ oil meal and peanut shell meal (Group B, n = 26). All animals received mineral supplements, active dry yeast, virginiamycin and monensin.

The experiment was conducted at the feedlot facility of “Embrapa Pecuária Sudeste” and lasted 105 days, which included 15 days for animal adaptation to the feedlot, 30 days for growth and 60 days for animal finishing. Feedlots were divided based on the dietary treatment and initial weights, with heavyweight and lightweight animals grouped separately (Table 1). The facility has collective stalls with automatic GrowSafe® (GrowSafe Systems Ltd., Airdrie, Alberta, Canada) feed system, used to collect data regarding live weight and daily food consumption. The CH<sub>4</sub> emission during the finishing period in the feedlot was measured using the GreenFeed system (C-lock Inc., Rapid City, SD, USA). All animal data used in this study is available in Table 1. Approximately 10 g of feces were obtained from each animal two weeks before slaughtering, and 50 mL of rumen content immediately after slaughter. All samples were frozen in liquid nitrogen and permanently stored at -80 °C prior to analysis. DNA extraction was performed using the Quick-DNA™ Fecal/Soil Microbe Miniprep Kit (ZYMO Research Corp., Irvine, CA), as determined by the standard protocol. PCR target amplification was performed using the follow primer sets: 341-b-S-17F (3'CCTACGGGNGGCWGCAG5') and 785-a-A-21R (3'GACTACHVGGGTATCTAATCC5') for bacteria 16S rRNA gene sequences; Ar915aF (3'AGGAATTGGCGGGGAGCAC5') and Ar1386R (3'GCGGTGTGTGCAAGGAGC5') for archaea 16S rRNA gene sequences; Reg1320R (3'AATTGCAAAGATCTATCCC5') and RP841F (3'GACTAGGGATTGGARTGG5') for protozoa 18S rRNA gene sequences. PCR conditions, sequencing libraries and DNA sequencing were performed as described in Andrade et al., (2020).

**Table 1 – Information regarding the initial weight group, slaughter date, dietary treatment and CH<sub>4</sub> dailymean of animals used in this study.**



Animal	Initial group	Dietary treatment	Slaughter date	CH <sub>4</sub> daily mean (g/d)	CH <sub>4</sub> (g/d) variation
238	Heavy	Group A	16/11/2016	173.37	56.98
239	Light	Group B	16/11/2016	186.64	33.74
240	Light	Group B	16/11/2016	157.54	77.17
242	Light	Group A	17/10/2016	178.71	34.68
244	Heavy	Group A	17/10/2016	168.35	42.17
246	Heavy	Group B	16/11/2016	161.67	63.95
464	Light	Group B	16/11/2016	170	N/A
466	Heavy	Group A	17/10/2016	135.07	45.94
468	Light	Group B	16/11/2016	164.52	47.46
470	Heavy	Group B	17/10/2016	154.59	75.06
474	Heavy	Group A	17/10/2016	210.34	66.97
479	Heavy	Group B	17/10/2016	137.3	39.74
482	Heavy	Group A	17/10/2016	201.23	46.69
483	Light	Group A	17/10/2016	199.54	52.91
490	Heavy	Group B	17/10/2016	161.25	58.3
491	Light	Group B	16/11/2016	162.3	59.75
492	Light	Group A	16/11/2016	201.69	46.6
494	Light	Group B	17/10/2016	194.75	38.54
499	Heavy	Group B	16/11/2016	160.67	52.6
500	Heavy	Group A	16/11/2016	210.24	77.43
502	Light	Group A	16/11/2016	149	N/A
505	Light	Group A	16/11/2016	164.89	48.58
506	Light	Group A	17/10/2016	176.82	54.84
507	Light	Group B	16/11/2016	217.03	60.53
510	Light	Group A	16/11/2016	181.87	55.56
511	Heavy	Group A	16/11/2016	210.48	70.68
514	Light	Group A	16/11/2016	191.81	56.02

515	Light	Group A	17/10/2016	206.19	53.95
516	Light	Group B	16/11/2016	150.51	37.45
517	Heavy	Group A	17/10/2016	156.89	32.16
520	Light	Group B	17/10/2016	201.08	34.17
521	Heavy	Group B	17/10/2016	152.2	68.05
523	Light	Group A	16/11/2016	176.48	40.73
1460	Light	Group A	16/11/2016	201.35	52.65
1462	Heavy	Group B	17/10/2016	120.64	39.55
1464	Heavy	Group B	17/10/2016	159.32	50.5
1468	Heavy	Group A	17/10/2016	145.93	47.5
1476	Heavy	Group A	17/10/2016	129.03	56.04
1479	Heavy	Group B	17/10/2016	169.32	54.2
1480	Light	Group A	16/11/2016	195.1	70.62
1481	Heavy	Group A	16/11/2016	176	56.32
1485	Heavy	Group B	17/10/2016	133.64	40.2
1493	Heavy	Group A	17/10/2016	141.27	42.87
1494	Light	Group B	17/10/2016	152	30.22
1495	Light	Group B	16/11/2016	193.65	49.55
1496	Heavy	Group B	17/10/2016	88.52	N/A
1498	Light	Group B	16/11/2016	183.68	71.71
1500	Heavy	Group A	17/10/2016	151.66	43.03
1501	Heavy	Group B	16/11/2016	156.4	44.07
1502	Light	Group A	16/11/2016	223.58	45.2
1503	Heavy	Group B	17/10/2016	157.81	72.5
1504	Light	Group B	16/11/2016	139.29	30.37

### Data retrieval, pre-processing and analysis

In addition to the dataset generated in this study, raw reads generated in our previous work with bulls fed conventional diet (Group A) were retrieved from the SRA database [accession number PRJNA525838], and processed to infer the impact of dietary treatments and to search for association with phenotypes.

All rRNA gene sequence reads from group A and B were filtered for quality (>Q25) and trimmed at the positions 220 (forward) and 175 (reverse) using QIIME 2 (version 2018.8) [29]. These positions were selected based on aggregation plots provided by QIIME2. The filtered data was submitted to DADA2 to generate Amplicon Sequence Variants (ASVs) with the option just-concatenate, and exclude chimeric sequences [30]. Bacterial sequences were classified using the SILVA database version 132 [31], archaeal sequences using the Rumen and Intestinal Methanogen database (RIM-DB) [32], which allows the classification of archaea up to the species level and protozoa using a curated database containing protozoa 18S rRNA genes sequences [33] with the feature classifier plugin within QIIME 2. Rarefaction curves were generated for each dataset and used to standardize the data (Additional file 1: Figure S1A, Additional file 2: Figure S2A). The resulting ASV table was used to determine alpha (Number of ASVs and Shannon-Wiener index) and beta diversities (Unweighted Unifrac distance) with QIIME 2.

## Statistical analysis

The Mixed models approach using the REML methodology was used in order to verify if the means of CH<sub>4</sub> emission among experimental groups was significantly different ( $P < 0.05$ ), using the dietary treatment as fixed effect and the slaughter batch and residual as random effects. We assessed differences in the microbial community structure, using alpha and beta diversities and the statistical tests Kruskal-Wallis, Permanova, and Principal Coordinates Analysis (PCoA) within QIIME 2 (version 2018.8) after data rarefaction. Relative abundances were transformed using the Centered Log-Ratio method (CLR), available in the python package scikit-bio (<http://scikit-bio.org/>) to be used in further analysis. We contrasted the microbiome of groups submitted to different dietary treatments using the Analysis of Composition of Microbiomes (ANCOM) V2.1 [34], with significance values adjusted for multiple tests using the Benjamin-Hochberg method ( $\alpha < 0.05$ ). We also applied a conservative W-statistic (W-statistic cutoff = 0.9), in which an ASV was considered as differential abundant if its composition varied when compared to 90% of the rest of the dataset, being the W-value the number of times the null hypothesis was rejected for a given ASV across two groups. ANCOM is a statistical approach that compares the abundances of each ASV individually transformed in Aitchison's log-ratio with all the remaining ASVs without any distributional assumptions.

We implemented Random Forest (RF) classification models to verify the use of microbiome population abundances as predictors to discriminate treatment groups. The models were trained using Python's scikit-learn package (<http://scikit-learn.org/>) with 70% of the data for the training set and 30% for the test set with five-fold cross-validation, max\_leaf\_nodes=30 and n\_estimators=100. Features and hyper-parameters were optimized using scikit-learn functions Feature Importance and GridSearch.

Association analysis with CH<sub>4</sub> emission phenotype was conducted using the Linear Mixed Models approach, similar to Difford et al., [35] with CLR-transformed abundances. The regression model was implemented using the Python's Statsmodel package (<http://statsmodel.org/>) using the following formula:

*CH<sub>4</sub> daily mean (g/d) ~ CLR(ASV) + Weight group + Slaughter date*

Experimental group information was used as covariates, diet as fixed effect and a random intercept for each group as a random effect in the statistical models. Significance values were adjusted for multiple tests using the Benjamin-Hochberg method with an exploratory significance level ( $\alpha \leq 0.1$ ). We considered only ASVs that presented a relative abundance higher than 0.5% and a prevalence in >10% of animals in experimental groups.

## Results

### Microbiome composition

The sequencing of microbiome rRNA amplicons from ruminal and fecal samples of the experimental group B yielded a total of 10,573,763 paired-end reads (4,628,604 paired-end reads for bacteria, 4,443,390 for archaea and 1,501,769 for protozoa), reaching 20,241,296 paired-end reads with the addition of sequencing data from Group A. After quality control, and singleton exclusion, a total of 4,519 bacterial ASVs (2,680 ruminal ASVs and 1,839 fecal ASVs), 1,023 archaeal ASVs (421 ruminal ASVs and 602 fecal ASVs) and 151 ruminal protozoan ASVs across treatments. Rarefaction curves based on the alpha-diversity metrics of Shannon-Wiener (diversity) reached a plateau, which indicated that additional sequences would not likely result in additional features.

Comparison of samples from different treatment groups using alpha-diversity metrics (Observed ASVs and Shannon-Wiener indexes) under the Kruskal-Wallis testing method revealed that rumen bacterial diversity was significantly more abundant and richer in animals fed the conventional diet (Group A) than those fed the byproducts diet (Group B) ( $P = 0.006$  and  $P = 0.04$ , respectively). Similarly, the ruminal archaea diversity was also richer ( $P = 0.0004$ ), but not more abundant. There was no significant difference when contrasting alpha diversity metrics of fecal samples. Comparisons of the beta-diversity metric Unweighted Unifrac using the PERMANOVA approach, revealed that samples of archaea and bacteria tended to form two significant clusters, which represented the treatment groups (adjusted  $P < 0.01$ ) (Supplementary Figures 1-3), a tendency most pronounced in fecal populations.

### Phenotypic description

Methane emission was calculated for each experimental group as the average value of all visits to the GrowSafe feedlot during the finishing period. Animals from the experimental group A presented a mean methane emission of 179,1 g/day and a standard error of 26,18 g/day, while animals from the experimental group B presented a mean methane emission of 161 g/day and a standard variation of 26,05 g/day. Mixed Models showed that the difference in CH<sub>4</sub> emission between experimental groups was significant ( $P < 0.0001$ ).

### Taxonomic composition of the experimental group B

Herein, we will present only results concerning the taxonomic composition of dietary treatment group B. Please refer to [28] for an extensive exploration of the taxonomic diversity of the group A.

The phylum Bacteroidetes was the most relatively abundant bacterial phylum identified in the rumen microbiome ( $38.18\% \pm 3.86\%$ ), followed by Firmicutes ( $35.72\%$ ), Proteobacteria ( $8.96\%$ ), Spirochaetes ( $5.40\%$ ) and Fibrobacteres ( $4.6\%$ ). Differently, the Phylum Firmicutes was the most abundant in the fecal microbiome ( $52.59\%$ ), followed by Bacteroidetes ( $30.87\%$ ), Proteobacteria ( $13.3\%$ ) and Tenericutes ( $1.31\%$ ). At the genus level, *Prevotella* was the most abundant genus in the rumen microbiome ( $19.87\%$ ), followed by *Treponema* ( $6.28\%$ ), *Ruminobacter* ( $5.78\%$ ), *Fibrobacter* ( $5.56\%$ ) and Christensenellaceae R-7 ( $5.56\%$ ). Conversely, the genus *Ruminococcaceae* UCG-005 was the most relatively abundant in the fecal microbiome ( $13.63\%$ ), followed by *Succinivibrio* ( $12.75\%$ ), *Bacteroides* ( $9.71\%$ ), *Prevotella* ( $6.69\%$ ) and Rikenellaceae RC9 ( $5.07\%$ ) (Figure 1).

Regarding the archaea domain, Euryarchaeota was the only phylum identified in both microbiomes. At the species level, these microbiomes were populated by *Methanobrevibacter gottschalkii* ( $59.16\%$  and  $74.89\%$  for rumen and feces, respectively), *Methanobrevibacter ruminantium* ( $31.98\%$  and  $17.11\%$ ) and *Methanosphaera* sp. ISO3-F5 ( $7.27\%$  and  $7.5\%$ ) (Figure 2A). As for protozoa, Ciliophora was the only phylum identified in rumen, and was populated by 3 genera, *Bozasella/Triplumaria* ( $70.73\%$ ), *Entodinium* ( $28.82\%$ ) and *Ostracodinium* ( $0.44\%$ ) (Figure 2B).

### Differential abundant ASVs in dietary treatment groups

We applied the analysis of composition of microbiomes (ANCOM) to investigate the influence of dietary treatments in the microbiome composition at the ASV level. Seventeen ruminal ASVs of bacterial origin were differentially more abundant (DA) in the group A, from which the most prominent were classified as Bacteroidales F082 group (ASV 20 and 23, CLR: 1.51), Christensenellaceae (ASV 112, CLR: 1.3), Pedosphaeraceae families (ASV 145, CLR: 1.09) and the genus Succiniclasticum (ASV 170, CLR: 1.04). Ten ASVs were DA in the group B, of which the most abundant were classified as Succiniclasticum (ASV 97, CLR: 0.48), Acetitomaculum (ASV 116, CLR: 1.07), Lachnospiraceae family (ASV 247, CLR: 0.98), Fibrobacter (ASV 96, CLR: 0.98) and Succinivibrio genus (ASV 118, CLR: 0.94) (Supplementary Figure 4). Also, three fecal ASVs were DA in our experimental groups; one was classified as a member of the family Rikenellaceae (ASV 361, CLR: 0.59) and was more abundant in the group A, while an ASV was classified as a member of the family Prevotellaceae (ASV 332, CLR: 0.51) and another as the genus Oscillibacter (ASV 526, CLR: 0.51) were both more abundant in the group B (Supplementary Figure 5).

Eight archaeal ASVs were DA among treatment groups in the rumen microbiome. Four ASVs classified as *M. gottschalkii* (ASVs 1, 2, 13 and 11, CLR > 1), one as *M. ruminantium* (ASV 23, CLR: 1.13) and one ASV belonging to the Methanomassiliicoccaceae family (ASV 36, CLR: 0.78) were all more abundant in the group A, while one classified as *M. ruminantium* (ASV 4, CLR: 1.79) and other as *Methanosphaera* group ISO3-F5 (ASV 33, CLR: 0.33) were more abundant in the group B (Supplementary Figure 6). Seven archaeal ASVs were DA in the fecal microbiome. From these, the ASVs classified as *M. gottschalkii* (ASVs 2, 13 and 11, CLR > 1.5) and *M. smithii* (ASV 28, CLR: 1.19) were more abundant in the group A, while *M.*

*ruminantium* (ASV 4, CLR: 2) and Methanosphaera group ISO3-F5 (ASVs 5 and 33, CLR > 0.8) were more abundant in the group B (Supplementary Figure 7). No DA ASVs of protozoa origin were observed.

### **Discrimination between dietary treatment groups with Random Forest classification models**

Random forest (RF) classification models were trained using CLR transformed relative abundances of each dataset, to test if the microbiome populations at ASV level could be used to discriminate the treatment group. Random forest has been shown to be the most accurate Machine Learning (ML) model for microbiome data analysis [36]. This method has the ability to discriminate groups, while considering interrelationships in high dimensional data [37]. The trained models resulted in high cross-validation scores for the bacteria test sets ( $r^2=0.89$  for rumen,  $r^2=0.84$  for feces), for archaea ( $r^2=0.86$  for rumen and  $r^2=0.82$  for feces) but not for protozoa ( $r^2=0.57$ ).

The feature importance function was used to select ASVs that contributed the most to the model's accuracy and to optimize the models. In short, the number of predictors were reduced to those with a contribution value  $\geq 0.01$  to retrain the models, this resulted in 16 of 1683 as predictors for bacteria, 27 of 118 for archaea, and 30 of 52 for protozoa in rumen, while 22 of 1077 ASVs were predictors for bacteria and 33 of 88 for archaea in feces, respectively. This feature reduction resulted in an increased cross-validation for bacteria ( $r^2=0.91$  for rumen and  $r^2=0.94$  for feces), archaea ( $r^2=0.91$  for rumen and  $r^2=0.86$  for feces) and for protozoa ( $r^2=0.71$ ) with high recall and precision scores (Supplementary Table 1). Predictors used are available in the Supplementary Table 2.

### **Association between bacterial and archaeal ASVs found in rumen and feces and CH<sub>4</sub> emission.**

Previous analysis showed a significant difference in the mean CH<sub>4</sub> emission of experimental groups, with group A (estimated mean = 179.11) emitting more methane than group B (estimated mean = 160.97). In order to investigate the proportion of variation of CH<sub>4</sub> emissions explained by the microbiome composition of these animals, Linear mixed models were used with experimental groups information as fixed effects, weight and slaughter groups as co-variables, daily mean CH<sub>4</sub> emission (g/day) as the dependent variable and individual log-transformed ASVs abundances as independent variables. This analysis identified significant associations between bacteria and archaea and CH<sub>4</sub> emission in both environments. Within the rumen microbiome, the ASV 40, a *Pseudobutyribrio* ( $\beta=16.5$ , contribution = 9.7%) and the ASV 44, a *Bacteroidales* ( $\beta=-2.6$ , contribution = -1.3%), were associated with CH<sub>4</sub> emission phenotype (Figure 3).

Furthermore, we identified two bacterial ASVs in the fecal microbiome that were positively associated with CH<sub>4</sub> emission: ASV 0, a *Succinivibrio* ( $\beta=10.2$ , contribution = 6%) and the ASV 36, a *Parabacteroides* ( $\beta=2.9$ , contribution = 1.7%). Also, there were four bacterial ASV in this biome that were negatively associated with CH<sub>4</sub> emission, ASV 35, a *Ruminococcaceae* UCG-005 ( $\beta=-12.5$ , contribution = -7.3%), ASV 39, a *Phascolarctobacterium* ( $\beta=-3.6$ , contribution = -1.8%), ASV 43, a *Bacteroides* ( $\beta=-2.9$ , contribution = -1.7%) and ASV 51, an *Akkermansia* ( $\beta=-2.5$ , contribution = -1.5%) (Figure 3). In addition, a single archaea

ASV classified as *M. gottschalkii* was identified as positively associated with CH<sub>4</sub> emission, the ARQ ASV 1 ( $\beta=4.21$ , contribution = 2.4%). There was no significant associations between CH<sub>4</sub> emission and fecal archaea ASVs and protozoa ASVs.

## Discussion

In our previous study, we extensively explored the taxonomic structure and relationships of bacteria, archaea and protozoa from two different sections of the Nelore cattle GIT [28]. Herein we expand this study by introducing a new experimental group under a different dietary treatment. We contrasted these experimental groups to investigate the impact of the dietary intervention on microbial abundance and diversity, as well as the impact of individual ASVs on host CH<sub>4</sub> emission.

### The microbiome structure is affected by the feed composition

Analysis with alpha-diversity metrics showed that both bacteria and archaea communities only differed in the rumen environment, being less rich in animals of the group B. Although being outside the scope of this study, a link between a poorer ruminal microbiome and the increase of the feed efficiency phenotype was detected, and evidences suggest that high efficient animals produces less methane [38,39]. As it will be further discussed, our methane association analysis reinforces the hypothesis of a favorable effect of the poorer microbiome on this trait. Also, PCoA analysis with the beta-diversity metric Unweighted Unifrac showed the existence of distinct clusters for treatment groups A and B for bacteria and archaea but not for protozoa. Altogether, these results indicate that feed is an important modulator of the microbiome, which agrees with previous studies in which the impact of different diets and feed components were evaluated [40,41].

Differential abundance analysis with bacterial ASVs revealed a significant impact of the dietary treatment in the bacterial populations of both environments. ASVs classified as belonging to the Christensenellaceae family, as well as to the *Prevotella*, and *Fibrobacter* genera, which are all producers of Short-Chain Fatty Acids (SCFA) such as acetate and butyrate [42,43], were identified as more abundant in the group A, while ASVs classified as genera known to produce succinate and propionate, *Succinoclasticum* and *Succinivibrio* as well as the Lachnospiraceae family [44,45], were identified as more abundant in group B. Differently from acetate and butyrate production, which increases H<sub>2</sub>, and consequently CH<sub>4</sub> production and emission [46], propionate is an electron acceptor end-product of rumen fermentation that is probably an alternative to methanogenesis [47]. Also, it was shown that the increase in propionate concentrations is strongly associated with a decrease in CH<sub>4</sub> production [48]. The three DA ASVs identified in the fecal samples corresponded to bacteria that commonly inhabit the hindgut, such as the *Oscillibacter* genus and Prevotellaceae family, both more abundant in the group B, and Rikenellaceae family, more abundant in the group A [49–51]. The identification of a small number of DA ASVs in the fecal microbiome is consistent with the alpha diversity analysis, in which there was no significant difference in both abundance and richness among experimental groups.

The dietary treatment also had a significant impact in the archaea populations that increased the abundance of ASVs classified as *M. gottschalkii* in both rumen and feces of animals from the group A, and *M. ruminantium* of animals from the group B. The relative abundance of these species differed in the experimental group under by-product diet (group B), in which it was observed a decreased relative abundance of *M. gottschalkii* and an increased abundance of *M. ruminantium* (9.6% and 23.8%) in rumen and feces when compared to conventional (group A). A study on sheep with contrasting phenotypes for CH<sub>4</sub> emission found a higher abundance of the archaea *M. gottschalkii* in the higher emitter group and *M. ruminantium* in the lower emitter group [10]. This difference can be explained by specific genomic structures because, unlike *M. gottschalkii*, *M. ruminantium* lacks the coding genes for methyl-CoM reductase II (McrII), affecting its fitness in an environment with high concentrations of H<sub>2</sub>, the main substrate for the hydrogenotrophic pathway of ruminal methanogenesis [52]. Furthermore, the relative abundance of these highly abundant ASVs can partially explain the difference in methane emission observed between treatment groups, in which the group B emitted less methane than group A.

We also built RF models to test if the microbial ASVs CLR-transformed abundances could be used as predictors for host's outcomes, in this specific case, the treatment group. Random Forests is a non-parametric ensemble machine learning approach, consisting in a collection of a multitude of decision trees, the forest, in which their predictions are averaged in a regression task, or selected based on a majority vote in a classification task. The R<sup>2</sup> score had a significant increase when re-trained with ASV that were contributing the most to the average reduction of weighted impurity in a tree, thus being more important for the classification. A small part of the bacterial ASVs identified as DA by the ANCOM approach, such as the ASVs 97 and 521 for rumen and ASVs 332 and 526 for feces, were selected for the re-training. On the other hand 12/15 of the archaea ASVs identified as DA were selected for re-training, which suggests that the archaeal populations are more sensitive to changes in the dietary treatments.

Random Forest models have been applied to the microbiome field to classify experimental groups based on the microbiome composition [36], to identify fecal contamination in environmental samples [53] and to identify taxa whose abundances were different in mothers delivering prematurely [54]. Altogether, these results indicate that the microbiome composition is affected by the feed at the individual ASV level.

### **Phenotypic associations indicates biomarkers for CH<sub>4</sub> emission in both sections of the gut**

Mixed Linear model analysis identified a single bacterial ASV, the ASV40, as positively associated with CH<sub>4</sub> emission in the rumen microbiome. This highly abundant ASV was classified as a member of the *Pseudobutyrvibrio* genus and presented the highest  $\beta$  coefficient, which explained 9.72% of the variation in CH<sub>4</sub> emission in the experimental groups. Butyrate-producing bacteria, such as *Pseudobutyrvibrio* and *Butyrvibrio*, were identified by Partial least squares as highly associated with CH<sub>4</sub> emissions in a study with *Bos taurus* breeds representing extreme phenotypes [55], thus confirming our findings with a different methodology.



One ASV classified as Bacteroidales F082 group, abundant in the rumen of different cattle breeds and ruminant species [56] was identified as negatively associated with CH<sub>4</sub> emission levels in the rumen microbiome. A positive association with CH<sub>4</sub> emission was described by Difford et al., (2018) for the uncultured Bacteroidales BS11 gut group. His claim was supported by the functional annotation of two genera inside this group, in which the end products for cellulose fermentation included acetate, butyrate, propionate, CO<sub>2</sub> and H<sub>2</sub> [35,57]. Although being commonly identified in the rumen microbiome of different species, the Bacteroidales order is genetically diverse, as one would expect of members of low taxonomic rank to be, thus reinforcing the need for additional exploration of this variability and for more studies in order to investigate the functional pathways responsible for the negative association of the F082 group and CH<sub>4</sub> emission.

Also, we identified 6 fecal ASVs that were significantly associated with CH<sub>4</sub> emission. This is the first time that such a relationship has been observed in beef cattle. Of the 6 ASVs, two were positively associated with this phenotype, a *Succinivibrio* and a *Parabacteroides*, which together explained 7.7% of the variance in CH<sub>4</sub> emission. The four negatively associated were *Akkermansia*, *Bacteroides*, *Phascolarctobacterium* and Ruminococcaceae UCG-005. Except for *Succinivibrio*, these genera were described as butyrate producers in the rumen biome [58–61], a short-chain fatty-acid that displays strong anti-inflammatory properties, modulates intestinal motility and improves the epithelial defense barrier in other mammals, such as mice and humans [62].

ASVs classified as Ruminococcaceae UCG-005, the most abundant genus in the fecal microbiome, one classified as *Phascolarctobacterium*, one classified as *Bacteroides* and one classified as *Akkermansia*, all propionate-producers, were negatively associated with CH<sub>4</sub> emission [63–65]. Propionate concentration showed a negative association with CH<sub>4</sub> production in the rumen environment, being considered an alternative [H] sink to methane [48]. Propionate is a short-chain fatty acid known to influence lipid biosynthesis, satiety, hunger, energy intake and even feeding behavior [66,67], traits that could affect the host metabolism and complex phenotypes, such as CH<sub>4</sub> emission. Although extensively studied in other model animals, the role of scfa-producing bacteria in the hindgut of bovines and the reasons behind these significant associations are yet to be understood. Further studies using more layers of information, such as meta-metabolomics and deep sequencing metagenomics, will be needed to investigate the metabolic background of these potential biomarkers. Regarding the archaea population, one ASVs classified as *M. gottschalkii* in the ruminal environment was positively associated with CH<sub>4</sub> emission. Members of *M. gottschalkii* species are the primary contributors to CH<sub>4</sub> production in the rumen microbiome [4]. Our results agrees with there previous findings, indicating a direct relationship between CH<sub>4</sub> production and *M. gottschalkii* abundance.

The GIT is a continuous and interconnected system, and as part of the digestive process, ruminants regurgitate digesta to chew partially digested material. Due to these characteristics, previous studies have suggested that the microbial populations of other sections of the GIT, such as both buccal and fecal environments, can be proxies for the rumen microbiome [28,68]. The findings of these studies support the

hypothesis of using these microorganisms as markers for the host's complex phenotypes, such as CH<sub>4</sub> emission. Several scientific studies and international consortia have been trying to find means to mitigate CH<sub>4</sub> emissions, through the use of feed formulations, feed additives and anti-methanogen vaccines [69]. Some authors consider this problem intractable because the ruminal microbiota can rapidly adapt to external interventions [70]. Additional experiments need to be performed to test the potential markers identified in this exploratory study. However, understanding the biology of specific microorganisms that contribute to complex phenotypes may help to develop successful interventions for methane mitigation in bovines.

## Conclusion

The feed composition induced significant differences in abundance and richness of ruminal and fecal microbial populations. The dietary treatment based on industrial byproducts applied to our experimental groups had an impact on the microbiome diversity of bacteria and archaea, but not on protozoa. Microbiome components (ASVs) of bacteria and archaea can be successfully used to predict the treatment group, thus giving support to the hypothesis that the feed intervention can modulate microbiome abundance and diversity. Microbiome components were associated with CH<sub>4</sub> emission in both ruminal and fecal microbiomes. While ruminal ASVs are expected to be directly associated to CH<sub>4</sub> production and emission, given that we monitored rumen CH<sub>4</sub> emission in the feedlot, the relation of fecal ASVs with this trait is unclear, although they can be biomarkers for CH<sub>4</sub> emission in an easier to access sample. Therefore, both ruminal and fecal ASVs can be used as biomarkers for methane production and emission.

## List Of Abbreviations

GIT - Gastrointestinal tract

ASV - Amplicon Sequence Variant

ML – Machine Learning

DA – Differentially abundant

SCFA – short-chain fatty-acid

RF – Random Forest

CLR – Centered log-ratio

ILR – Isometric log-ratio

ALR – Additive log-ratio

# Declarations

## Ethics approval and consent to participate

Experimental procedures were conducted following Brazilian guidelines on animal welfare and approved by the Ethics Committee on the Use of Animals, College of Veterinary and Animal Science, São Paulo State University under protocol n° 8510190118.

## Availability of data and material

All sequencing data are available in the NCBI Sequence Read Archive (SRA), under the bioproject number PRJNA638250.

## Competing interests

All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

## Funding

São Paulo Research Foundation (FAPESP – 2019/1557-7)

National Council for Scientific and Technological Development (CNPq - 4281153-6/2018)

## Acknowledgments

We thank Dr. Marcela Morelli for collecting sample data, Dr. Leandro Sannomiya Sakamoto and the staff of the Embrapa Southeast Cattle Animal biotechnology laboratory for assistance on sample collection.

## Consent for publication

Not applicable

## Authors contributions

BGNA, JEK, AB, JCPP, JMR, LLC and LCAR conceived the experiment; BGNA and FAB performed the experiments; BGNA, RRCC, HA and GBM performed analysis; BGNA, HA, RRCC, PSNO, AZN, SRM and LCAR and MMS interpreted the results; BGNA and LCAR drafted and revised the manuscript. All authors read and approved the final manuscript.

## Funding

This study was conducted with funding from EMBRAPA, São Paulo Research Foundation scholarship to BGNA (FAPESP, grant number: 2019/15557-7), São Paulo Research Foundation (FAPESP, grant number: 2019/04089-2), the National Council for Scientific and Technological Development (CNPq, grant number: 428153/2018) and fellowships to LCAR and LLC.

# References

1. Cook J, Oreskes N, Doran PT, Anderegg WRL, Verheggen B, Maibach EW, et al. Consensus on consensus: A synthesis of consensus estimates on human-caused global warming. *Environ Res Lett.* IOP Publishing; 2016;11.
2. Williams M, Zalasiewicz J, Haff P, Schwägerl C, Barnosky AD, Ellis EC. The Anthropocene biosphere. *Anthr Rev.* 2015;2:196–219.
3. Smith P, Gregory PJ. Climate change and sustainable food production. *Proc Nutr Soc.* 2013;72:21–8.
4. Tapio I, Snelling TJ, Strozzi F, Wallace RJ. The ruminal microbiome associated with methane emissions from ruminant livestock. *J Anim Sci Biotechnol. Journal of Animal Science and Biotechnology*; 2017;8:1–11.
5. Tizioto PC, Decker JE, Taylor JF, Schnabel RD, Mudadu M a, Silva FL, et al. Genome scan for meat quality traits in Nelore beef cattle. *Physiol Genomics.* 2013;45:1012–20.
6. Henrique V, Correia L, Regitano DA, Geistlinger L, Morosini S, Zimmer R, et al. Genome-Wide Detection of CNVs and Their Association with Meat Tenderness in Nelore Cattle. 2016;1–22.
7. Mudadu MA, Porto-Neto LR, Mokry FB, Tizioto PC, Oliveira PSN, Tullio RR, et al. Genomic structure and marker-derived gene networks for growth and meat quality traits of Brazilian Nelore beef cattle. *BMC Genomics.* BMC Genomics; 2016;17:235.
8. Pszczola M, Strabel T, Mucha S, Sell-Kubiak E. Genome-wide association identifies methane production level relation to genetic control of digestive tract development in dairy cows. *Sci Rep.* 2018;8:15164.
9. Leahy SC, Kelly WJ, Ronimus RS, Wedlock N, Altermann E, Attwood GT. Genome sequencing of rumen bacteria and archaea and its application to methane mitigation strategies. *Animal.* 2013;7 Suppl 2:235–43.
10. Shi W, Moon CD, Leahy SC, Kang D, Froula J, Kittelmann S, et al. Methane yield phenotypes linked to differential gene expression in the sheep rumen microbiome. 2014;1517–25.
11. Flint HJ, Duncan SH, Scott KP, Louis P. Interactions and competition within the microbial community of the human colon: links between diet and health. *Environ Microbiol.* 2007;9:1101–11.
12. Belkaid Y, Hand TW. Role of the Microbiota in Immunity and Inflammation. *Cell.* 2014;157:121–41.
13. Goodrich JK, Davenport ER, Clark AG, Ley RE. The Relationship Between the Human Genome and Microbiome Comes into View. *Annu Rev Genet.* 2017;51:413–33.
14. Solden L, Lloyd K, Wrighton K. The bright side of microbial dark matter: lessons learned from the uncultivated majority. *Curr Opin Microbiol.* 2016;31:217–26.
15. Gilbert J a., Dupont CL. Microbial Metagenomics: Beyond the Genome. *Ann Rev Mar Sci.* 2011;3:347–71.
16. Vallianou NG, Stratigou T, Tsagarakis S. Microbiome and diabetes: Where are we now? *Diabetes Res Clin Pract.* 2018;146:111–8.

17. Wong SH, Yu J. Gut microbiota in colorectal cancer: mechanisms of action and clinical applications. *Nat Rev Gastroenterol Hepatol*. 2019;16:690–704.
18. Magro DO, Santos A, Guadagnini D, de Godoy FM, Silva SHM, Lemos WJF, et al. Remission in Crohn's disease is accompanied by alterations in the gut microbiota and mucins production. *Sci Rep*. 2019;9:13263.
19. Chang H-X, Haudenschild JS, Bowen CR, Hartman GL. Metagenome-Wide Association Study and Machine Learning Prediction of Bulk Soil Microbiome and Crop Productivity. *Front Microbiol*. 2017;8:519.
20. Canibe N, Højberg O, Højsgaard S, Jensen BB. Feed physical form and formic acid addition to the feed affect the gastrointestinal ecology and growth performance of growing pigs. *J Anim Sci*. 2005;83:1287–302.
21. Jami E, White BA, Mizrahi I. Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. *PLoS One*. 2014;9:e85423.
22. Noel SJ, Olijhoek DW, Mclean F, Løvendahl P, Lund P, Højberg O. Rumen and Fecal Microbial Community Structure of Holstein and Jersey Dairy Cows as Affected by Breed, Diet, and Residual Feed Intake. *Animals*. 2019;9:498.
23. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front Microbiol*. 2017;8.
24. Praus P. Robust multivariate analysis of compositional data of treated wastewaters. *Environ Earth Sci*. 2019;78:248.
25. van den Boogaart KG, Tolosana-Delgado R. Analyzing Compositional Data with R. *Anal Compos Data with R*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 95–175.
26. de Oliveira MNV, Jewell KA, Freitas FS, Benjamin LA, Tótola MR, Borges AC, et al. Characterizing the microbiota across the gastrointestinal tract of a Brazilian Nelore steer. *Vet Microbiol*. 2013;164:307–14.
27. Granja-Salcedo YT, Fernandes RM, de Araujo RC, Kishi LT, Berchielli TT, de Resende FD, et al. Long-Term Encapsulated Nitrate Supplementation Modulates Rumen Microbial Diversity and Rumen Fermentation to Reduce Methane Emission in Grazing Steers. *Front Microbiol*. 2019;10:614.
28. Andrade BGN, Bressani FA, Cuadrat RRC, Tizioto PC, Oliveira PSN De, Mourão GB, et al. The structure of microbial populations in Nelore GIT reveals inter-dependency of methanogens in feces and rumen. *Journal of Animal Science and Biotechnology*; 2020;1–10.
29. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet C, Al-Ghalith GA, et al. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ*. 2018;9–10.
30. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13:581–3.
31. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41:590–6.

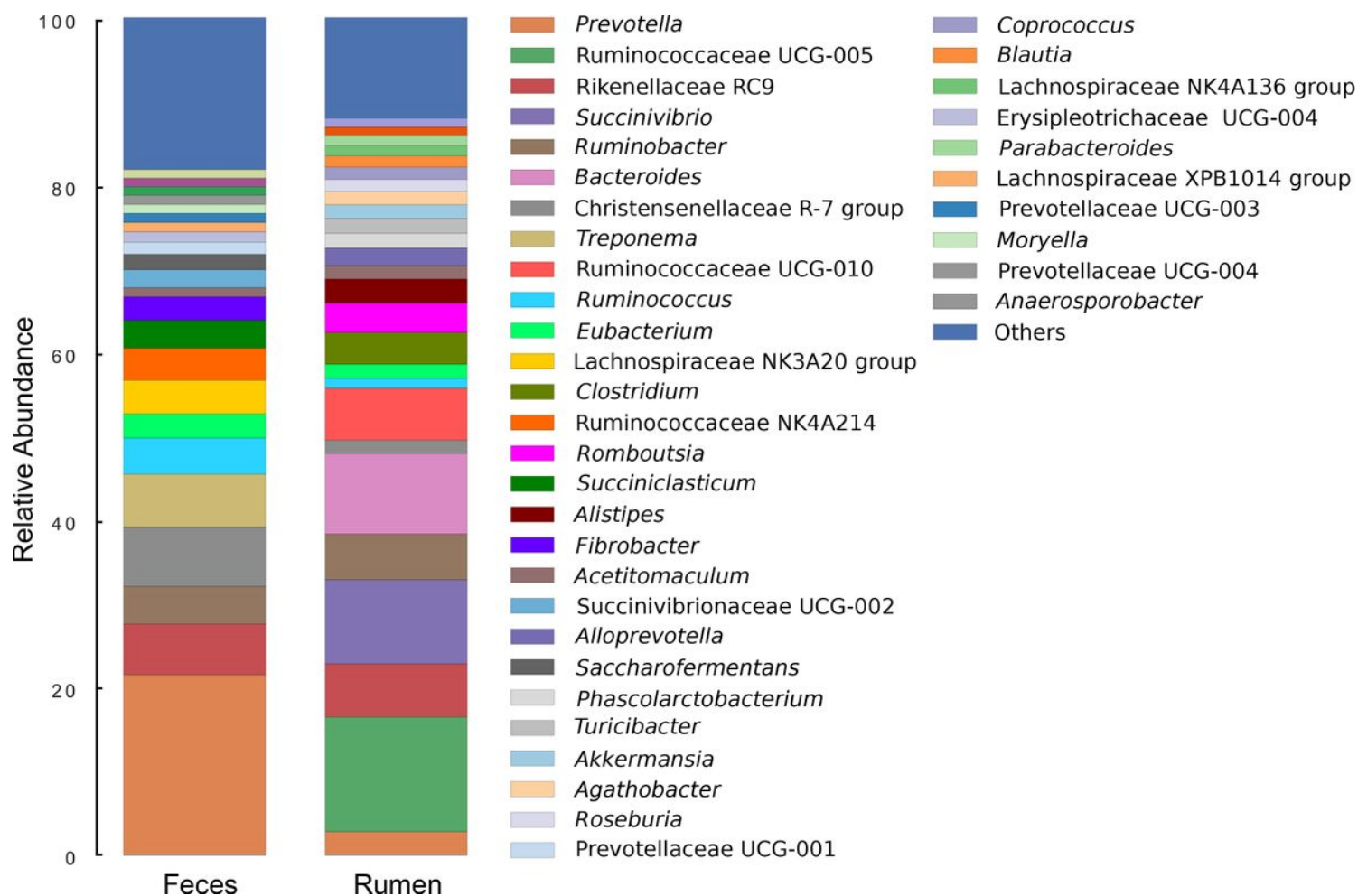
32. Seedorf H, Kittelmann S, Henderson G, Janssen PH. RIM-DB: a taxonomic framework for community structure analysis of methanogenic archaea from the rumen and other intestinal environments. *PeerJ*. 2014;2:e494.
33. Kittelmann S, Devente SR, Kirk MR, Seedorf H, Dehority BA, Janssen PH. Phylogeny of Intestinal Ciliates, Including *Charonina ventriculi*, and Comparison of Microscopy and 18S rRNA Gene Pyrosequencing for Rumen Ciliate Community Structure Analysis. *Appl Environ Microbiol*. 2015;81:2433–44.
34. Kaul A, Mandal S, Davidov O, Peddada SD. Analysis of microbiome data in the presence of excess zeros. *Front Microbiol*. 2017;8:1–10.
35. Difford GF, Plichta DR, Løvendahl P, Lassen J, Noel SJ, Højberg O, et al. Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. *PLoS Genet*. 2018;14:1–22.
36. Zhou YH, Gallins P. A review and tutorial of machine learning methods for microbiome host trait prediction. *Front Genet*. 2019;10:1–14.
37. Žižka J, Dařena F, Svoboda A. Random Forest. *Text Min with Mach Learn*. First. | Boca Raton: CRC Press, 2019.: CRC Press; 2019. p. 193–200.
38. Kruger Ben Shabat S, Sasson G, Doron-Faigenboim A, Durman T, Yaacoby S, Berg Miller ME, et al. Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. *ISME J*. Nature Publishing Group; 2016;10:2958–72.
39. Hegarty RS, Goopy JP, Herd RM, McCorkell B. Cattle selected for lower residual feed intake have reduced daily methane production. *J Anim Sci*. 2007;85:1479–86.
40. Popova M, Guyader J, Silberberg M, Seradj AR, Saro C, Bernard A, et al. Changes in the Rumen Microbiota of Cows in Response to Dietary Supplementation with Nitrate, Linseed, and Saponin Alone or in Combination. *Appl Environ Microbiol*. 2019;85:1–16.
41. Dill-McFarland KA, Weimer PJ, Breaker JD, Suen G. Diet Influences Early Microbiota Development in Dairy Calves without Long-Term Impacts on Milk Production. *Appl Environ Microbiol*. 2019;85:1–12.
42. Waters JL, Ley RE. The human gut bacteria Christensenellaceae are widespread, heritable, and associated with health. *BMC Biol*. BMC Biology; 2019;17:1–11.
43. Neumann AP, McCormick CA, Suen G. Fibrobacter communities in the gastrointestinal tracts of diverse hindgut-fermenting herbivores are distinct from those of the rumen. *Environ Microbiol*. 2017;19:3768–83.
44. Hespell RB. The Genera *Succinivibrio* and *Succinimonas*. *The Prokaryotes* [Internet]. New York, NY: Springer New York; 1992. p. 3979–82. Available from: [http://link.springer.com/10.1007/978-1-4757-2191-1\\_60](http://link.springer.com/10.1007/978-1-4757-2191-1_60)
45. van GYLWYK NO. *Succiniclasticum ruminis* gen. nov., sp. nov., a Ruminant Bacterium Converting Succinate to Propionate as the Sole Energy-Yielding Mechanism. *Int J Syst Bacteriol*. 1995;45:297–300.
46. Wolin MJ. A Theoretical Rumen Fermentation Balance. *J Dairy Sci*. 1960;43:1452–9.

47. Wang K, Nan X, Chu K, Tong J, Yang L, Zheng S, et al. Shifts of hydrogen metabolism from methanogenesis to propionate production in response to Replacement of forage fiber with non-forage fiber sources in diets in vitro. *Front Microbiol.* 2018;9:1–12.
48. Ungerfeld EM. Shifts in metabolic hydrogen sinks in the methanogenesis-inhibited ruminal fermentation: a meta-analysis. *Front Microbiol.* 2015;6.
49. Lee G-H, Kumar S, Lee J-H, Chang D-H, Kim D-S, Choi S-H, et al. Genome Sequence of *Oscillibacter ruminantium* Strain GH1, Isolated from Rumen of Korean Native Cattle. *J Bacteriol.* 2012;194:6362–6362.
50. Marounek M, Duskova D. Metabolism of pectin in rumen bacteria *Butyrivibrio fibrisolvens* and *Prevotella ruminicola*. *Lett Appl Microbiol.* 1999;29:429–33.
51. Chen X, Zuo Q, Hai Y, Sun XJ. Lactulose: An indirect antioxidant ameliorating inflammatory bowel disease by increasing hydrogen production. *Med Hypotheses.* 2011;76:325–7.
52. Leahy SC, Kelly WJ, Altermann E, Ronimus RS, Yeoman CJ, Pacheco DM, et al. The genome sequence of the rumen methanogen *Methanobrevibacter ruminantium* reveals new possibilities for controlling ruminant methane emissions. *PLoS One.* 2010;5:e8926.
53. Roguet A, Eren AM, Newton RJ, McLellan SL. Fecal source identification using random forest. *Microbiome.* 2018;6:1–15.
54. Dahl C, Stanislawski M, Iszatt N, Mandal S, Lozupone C, Clemente JC, et al. Gut microbiome of mothers delivering prematurely shows reduced diversity and lower relative abundance of *Bifidobacterium* and *Streptococcus*. *PLoS One.* 2017;12:1–14.
55. Auffret MD, Stewart R, Dewhurst RJ, Duthie CA, Rooke JA, Wallace RJ, et al. Identification, comparison, and validation of robust rumen microbial biomarkers for methane emissions using diverse *Bos Taurus* breeds and basal diets. *Front Microbiol.* 2018;8:1–15.
56. Xin J, Chai Z, Zhang C, Zhang Q, Zhu Y, Cao H, et al. Comparing the Microbial Community in Four Stomach of Dairy Cattle, Yellow Cattle and Three Yak Herds in Qinghai-Tibetan Plateau. *Front Microbiol.* 2019;10.
57. Danielsson R, Dicksved J, Sun L, Gonda H, Müller B, Schnürer A, et al. Methane production in dairy cows correlates with rumen methanogenic and bacterial community structure. *Front Microbiol.* 2017;8:1–15.
58. Vital M, Karch A, Pieper DH. Colonic Butyrate-Producing Communities in Humans: an Overview Using Omics Data. *mSystems.* 2017;2:1–18.
59. Wang K, Liao M, Zhou N, Bao L, Ma K, Zheng Z, et al. *Parabacteroides distasonis* Alleviates Obesity and Metabolic Dysfunctions via Production of Succinate and Secondary Bile Acids. *Cell Rep.* 2019;26:222-235.e5.
60. García-Mantrana I, Calatayud M, Romo-vaquero M, Espín JC, Selva M V, Collado MC. Urolithin Metabotypes Can Determine the Modulation of Gut Microbiota in Healthy Individuals. *Nutrients.* 2019;11:1–13.

61. El Hage R, Hernandez-Sanabria E, Calatayud Arroyo M, Props R, Van De Wiele T. Propionate-producing consortium restores antibiotic-induced dysbiosis in a dynamic in vitro model of the human intestinal microbial ecosystem. *Front Microbiol.* 2019;10:1–17.
62. Canani RB. Potential beneficial effects of butyrate in intestinal and extraintestinal diseases. *World J Gastroenterol.* 2011;17:1519.
63. Wu F, Guo X, Zhang J, Zhang M, Ou Z, Peng Y. *Phascolarctobacterium faecium* abundant colonization in human gastrointestinal tract. *Exp Ther Med.* 2017;14:3122–6.
64. Ouwerkerk JP, van der Ark KCH, Davids M, Claassens NJ, Finestra TR, de Vos WM, et al. Adaptation of *Akkermansia muciniphila* to the Oxic-Anoxic Interface of the Mucus Layer. Schloss PD, editor. *Appl Environ Microbiol.* 2016;82:6983–93.
65. Shimizu J, Kubota T, Takada E, Takai K, Fujiwara N, Arimitsu N, et al. Propionate-producing bacteria in the intestine may associate with skewed responses of IL10-producing regulatory T cells in patients with relapsing polychondritis. Wilson BA, editor. *PLoS One.* 2018;13:e0203657.
66. Oba M, Allen MS. Intraruminal Infusion of Propionate Alters Feeding Behavior and Decreases Energy Intake of Lactating Dairy Cows. *J Nutr.* 2003;133:c–1099.
67. Anderson JW, Chen WJL. Plant fiber. Carbohydrate and lipid metabolism. *Am J Clin Nutr.* 1979;32:346–63.
68. Tapio I, Shingfield KJ, McKain N, Bonin A, Fischer D, Bayat AR, et al. Oral Samples as Non-Invasive Proxies for Assessing the Composition of the Rumen Microbial Community. *PLoS One.* 2016;11:e0151220.
69. Moss AR, Jouany JP, Newbold J. Methane production by ruminants: Its contribution to global warming. *Anim Res.* 2000;49:231–53.
70. Roehe R, Dewhurst RJ, Duthie CA, Rooke JA, McKain N, Ross DW, et al. Bovine Host Genetic Variation Influences Rumen Microbial Methane Production with Best Selection Criterion for Low Methane Emitting and Efficiently Feed Converting Hosts Based on Metagenomic Gene Abundance. *PLoS Genet.* 2016;12:1–20.

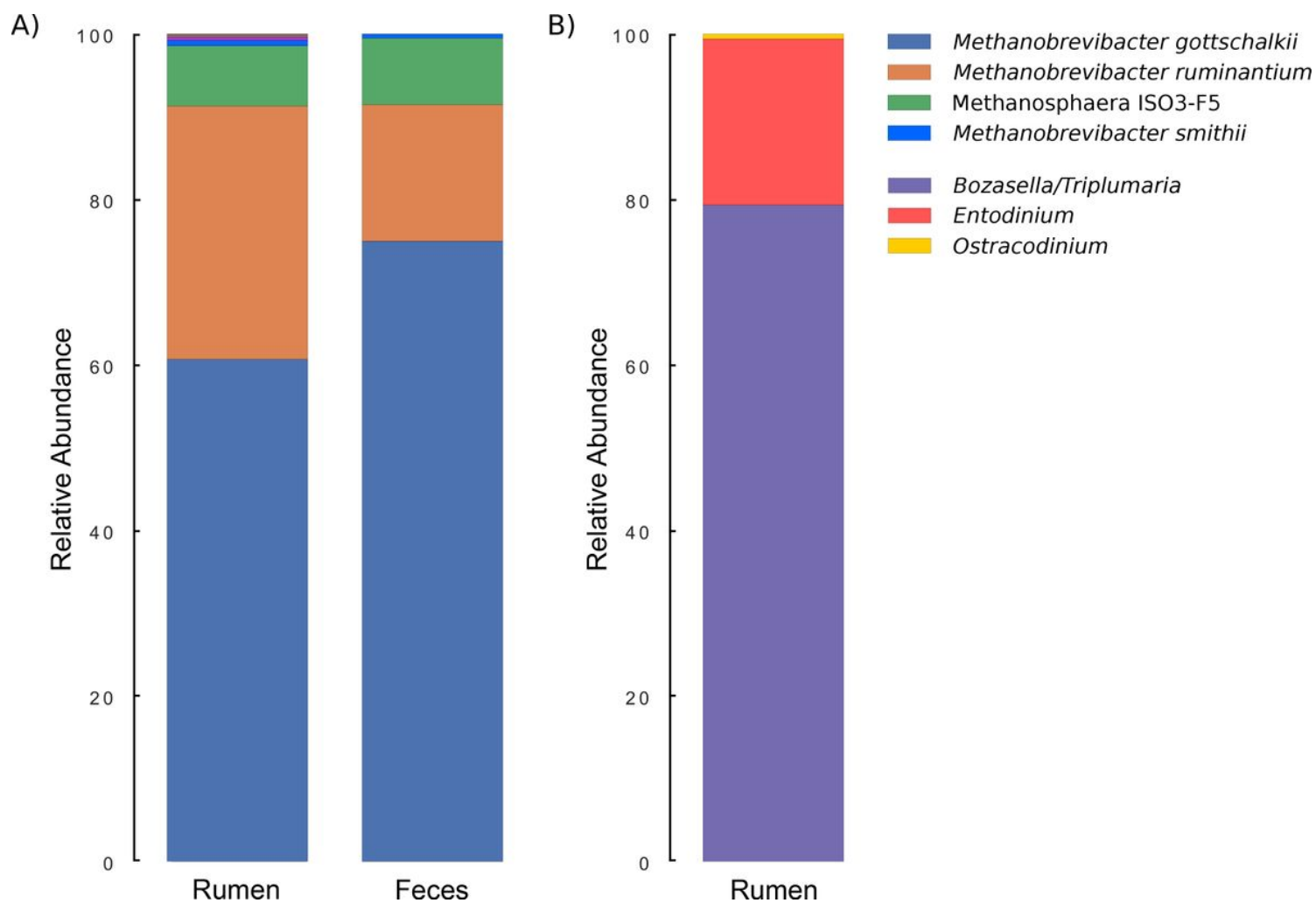
## Figures





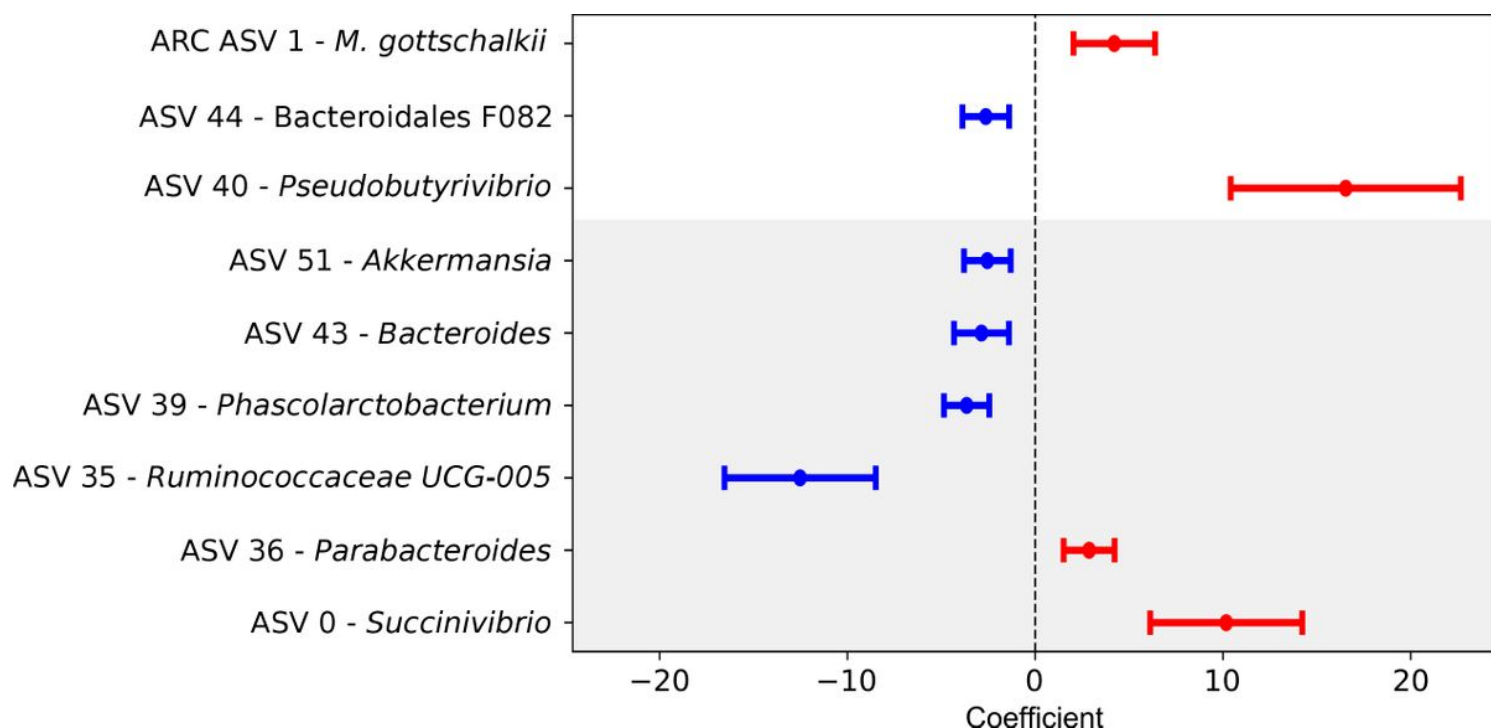
**Figure 1**

Relative abundance of bacterial populations, at genus level, in the rumen and feces of Nelore cattle fed with byproduct based diet. Only microorganisms with a relative abundance greater than 0.5% are shown in the legend.



**Figure 2**

Relative abundance of A) archaeal populations, at species level, in the rumen and feces of Nelore cattle fed with byproduct based diet and B) protozoa populations, at genus level, in the rumen of Nelore Cattle. Only microorganisms with a relative abundance greater than 0.5% are shown in the legend.



**Figure 3**

Errorbar plot showing the results of association analysis of ASVs from bacterial and archaea origins in rumen (in white) and fecal (in grey) samples with CH<sub>4</sub> emission phenotype. Blue bars represent negative associations while red bars represent positive associations.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigure7.pdf](#)
- [SupplementaryFigure6.pdf](#)
- [SupplementaryData.docx](#)
- [SupplementaryFigure5.pdf](#)
- [SupplementaryFigure4.pdf](#)
- [SupplementaryFigure3.pdf](#)
- [SupplementaryFigure2.pdf](#)
- [SupplementaryFigure1.pdf](#)